UDC 551.501.45:551.524.32:551.583.15

Time Series Analysis of Annual Temperatures

ERAT S. JOSEPH—*Lafayette College, Easton, Pa.*

ABSTRACT—The time series of annual mean, maximum, and minimum temperatures at Boulder, Denver, Fort Collins, and Pueblo, Colo., were analyzed to determine whether or not they are randomly distributed. Tests were performed by means of autocorrelation function and power spectrum. For series with significant first-order serial correlation coefficient, the adequacy of a first-order Markov process was investigated. The analysis revealed that the series of annual maximum temperatures at Fort Collins, and those of annual minimum temperatures at Fort Collins and Pueblo are purely random (white noise). For the remaining nine series, the values of autocorrelation function r_1 are significant, and a first-order Markov model seems adequate.

1. INTRODUCTION

A time series consists of observations arranged sequentially with respect to time. Yule (1921) and others noted that the time series of many natural phenomena belong to periodic, moving average, or autoregressive processes. Landsberg et al. (1959), Polowchak and Panofsky (1968), and others applied spectrum theory to the time series of daily, weekly, and monthly temperatures.

A time series is said to be randomly distributed if each event is statistically independent of all preceding and succeeding events. Many hydrometeorological processes are characterized by dependent events. Though dependence is sometimes considered a nuisance, its analysis often yields fruitful results.

2. TEST A: ANALYSIS OF AUTOCORRELATION FUNCTION

The autocorrelation function is the ratio of the autocovariance to the variance. In a time series with n data points, it may be estimated by

$$r_k=\left[\frac{1}{n-k}\sum_{i=1}^{n-k}x_i x_{i+k}-\frac{1}{(n-k)^2}\left(\sum_{i=1}^{n-k}x_i\right)\left(\sum_{i=1}^{n-k}x_{i+k}\right)\right]$$
$$\times\left\{\left[\frac{1}{n-k}\sum_{i=1}^{n-k}x_i^2-\frac{1}{(n-k)^2}\left(\sum_{i=1}^{n-k}x_i\right)^2\right]^{1/2}\right.$$
$$\left.\times\left[\frac{1}{n-k}\sum_{i=1}^{n-k}x_{i+k}^2-\frac{1}{(n-k)^2}\left(\sum_{i=1}^{n-k}x_{i+k}\right)^2\right]^{1/2}\right\}^{-1} \tag{1}$$

where r_k stands for the autocorrelation function of order k. The first-order autocorrelation function r_1 can be used to test whether the time series is purely random (white noise). Anderson (1941) developed confidence limits (C.L.) for r_1 on the basis of a circularly defined autocorrelation. He showed that, for a random normal series, r_1 is approximately normally distributed with mean $(-1)/(n-1)$ and variance $(n-2)/(n-1)^2$, and, hence,

$$\text{C.L. }[r_1]=\frac{-1\pm z_\alpha\sqrt{n-2}}{n-1} \tag{2}$$

where z_α is the standard normal variate corresponding to significance level α. When n is large, the probability distribution of r_1 can be taken to be normal with mean zero and variance $1/n$. If the estimated value of r_1 falls outside the confidence limits of equation 2, then r_1 is treated as significantly different from zero. Equation 2 may also be used to test the significance of r_k for $k>1$ if k is small relative to n (Matalas 1966).

A plot of r_k against k is called a correlogram. The shape of a correlogram, in principle, reveals the nature of a process. For a periodic process, the correlogram is periodic; for a moving average process, it vanishes; and for an autoregressive process, it is exponential. For a first-order Markov process, $r_k=r_1^k$ and the correlogram "decays" exponentially. Usually, when the number of data points is small, the correlogram fails to damp as expected because the observed autocorrelation functions are subject to inflation due to sampling errors.

3. TEST B: ANALYSIS OF POWER SPECTRUM

Power spectrum is the Fourier transform of the autocovariance function. Estimates of spectral densities can be used to test whether or not an observed time series could be regarded as the realization of a certain process.

The raw estimates of the power spectrum for maximum lag, m, are computed by

$$L_p=\frac{1}{\pi}\left[W_0+2\sum_{q=1}^{m-1}W_q\cos\left(\frac{\pi pq}{m}\right)+W_m\cos(\pi p)\right] \tag{3}$$

where w_p is the autocovariance function at lag p. The raw spectrum is smoothed to yield

$$U_p=0.23L_{p-1}+0.54L_p+0.23L_{p+1}. \tag{4}$$

TABLE 1.—*Significance of first-order serial correlation coefficients*

Station	Length of record, n	Estimated r_1 for indicated series			90 percent C.L.[r_1]	
		Annual mean temperature	Annual maximum temperature	Annual minimum temperature		
Boulder	73	0. 393	0. 387	0. 196	−0. 207	0. 179
Denver	99	. 194	. 305	. 392	− . 176	. 155
Fort Collins	82	. 329	. 136	− . 002	− . 194	. 169
Pueblo	82	. 367	. 226	. 023	− . 194	. 169

To determine whether the spectrum represents a certain process, we fit a null continuum (i.e., the spectrum according to a null hypothesis) and compare its local value to the spectral estimates at that wavelength.

Usually, when the first-order autocorrelation function is not significantly different from zero, the null hypothesis is a white noise process. The spectrum of a purely random time series is a horizontal straight line with amplitude the same everywhere. Thus, the appropriate null continuum for a white noise process is a horizontal straight line with amplitude equal to the mean of the spectral estimates.

If r_1 is significant, the adequacy of an autoregressive process may be tested. For a first-order Markov process, the null continuum, U^*, is approximately given by

$$U_p^*=C_p\overline{U}, \qquad 0\leq p\leq m \tag{5}$$

where $\overline{U}$ is the mean of the spectral estimates and

$$C_p=\frac{1-r_1^2}{1+r_1^2-2r_1 \cos \dfrac{\pi p}{m}}. \tag{6}$$

The confidence band for the spectrum of an assumed process is obtained by multiplying values of the null continuum by $\chi^2_\alpha(\nu)/\nu$ and $\chi^2_{1-\alpha}(\nu)/\nu$, where ν is the equivalent degree of freedom. Here, $\nu=2n/m$; n is the number of data points, and m is the maximum number of lags. (For detailed information see Jenkins 1961, Granger and Hatanaka 1964, and Mitchell 1966.) If none of the spectral estimates departs significantly from the null continuum; that is, if none falls outside the confidence band, the hypothesis of the adequacy of the assumed model is accepted.

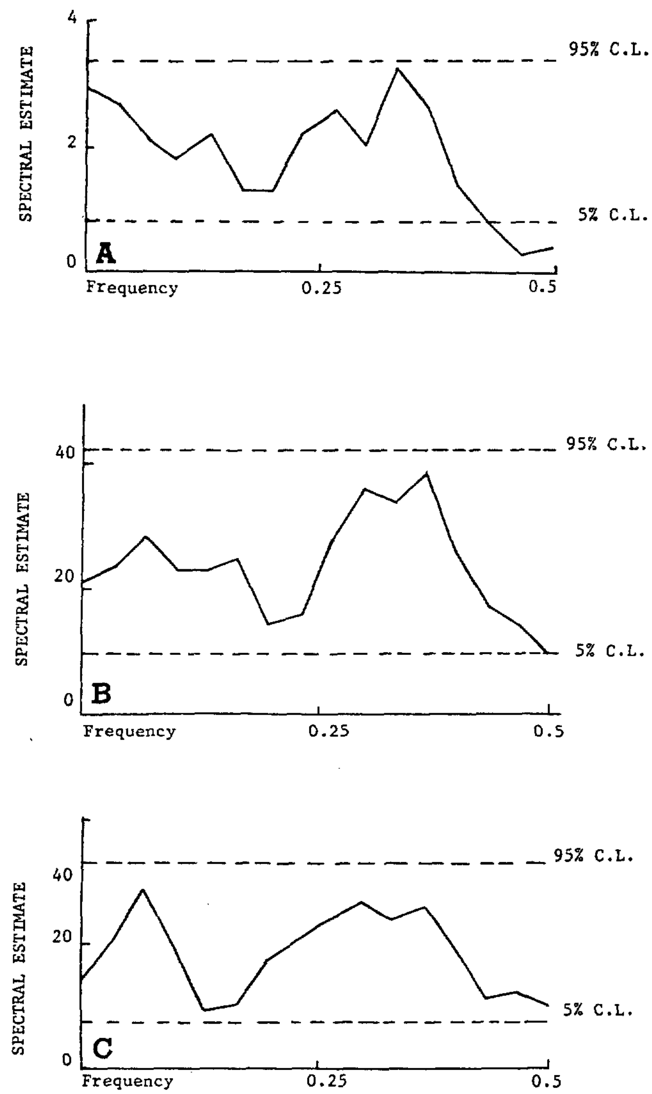

FIGURE 1.—White noise continuum with confidence limits and spectra of (A) annual maximum temperature at Fort Collins, (B) annual minimum temperature at Fort Collins, and (C) annual minimum temperature at Pueblo.

4. RESULTS OF ANALYSIS

The time series of annual mean, maximum, and minimum temperatures at Boulder, Denver, Fort Collins and Pueblo, Colo., were analyzed. These stations have relatively long records.

For the 12 series, the estimated values of the first-order autocorrelation functions with 90-percent confidence limits are tabulated in table 1. Based on the significance of r_1, the hypothesis of randomness was rejected by all except three series. Tests for the white noise process were performed on these three series using power spectrum and are graphically illustrated in figure 1. Those series with significant r_1 were tested for a first-order Markov process. Plots of estimated spectral densities for a minimum lag 15 with the confidence limits of the red noise continuum are shown in figure 2.

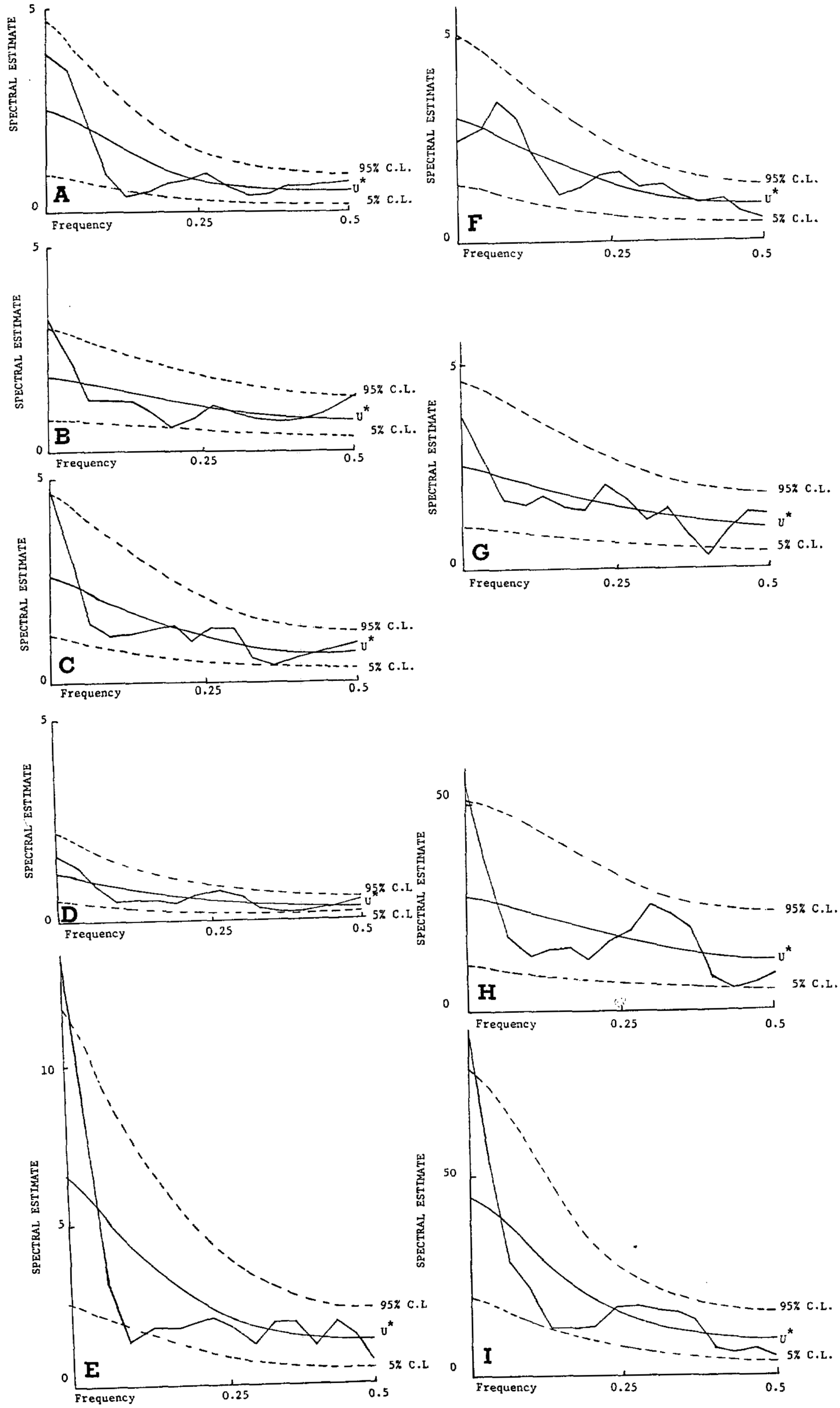

FIGURE 2.—Markov red noise continuum with confidence limits and spectra of annual mean temperature at (A) Boulder, (B) Denver, (C) Fort Collins, and (D) Pueblo; spectra of annual maximum temperature at (E) Boulder, (F) Denver, and (G) Pueblo; and spectra of annual minimum temperature at (H) Boulder and (I) Denver.

5. CONCLUSION

The investigation revealed that the series of annual maximum temperatures at Fort Collins and those of annual minimum temperatures at Fort Collins and Pueblo are purely random (white noise). For the remaining nine series, the values of autocorrelation function r_1 are significant, and a first-order Markov model seems adequate. The available data points in the series are not long enough to distinguish between apparent time trends and chance events.

ACKNOWLEDGMENTS

The author is indebted to Roger G. Barry, University of Colorado, for giving his valuable suggestions and to Jack D. Ives for making available the facilities at the Institute of Arctic and Alpine Research (INSTAAR). Financial support for this research was provided by the National Science Foundation under the Research Participation Program for College Teachers—1971 at INSTAAR, University of Colorado.

REFERENCES

Anderson, R. L., "Distribution of Serial Correlation Coefficient," *Annals of Mathematical Statistics*, Vol. 8, No. 1, Mar. **1941**, pp. 1–13.

Granger, C. W. J., and Hatanaka, M., *Spectral Analysis of Economic Time Series*, Princeton University Press, New Jersey, **1964**, 299 pp. (See pp. 1–262.)

Jenkins, Gwilym M., "General Considerations in Analysis of Spectra," *Technometrics*, Vol. 3, No. 2, **1961**, pp. 133–136.

Landsberg, Helmut E., Mitchell, J. Murray Jr., and Crutcher, Harold L., "Power Spectrum Analysis of Climatological Data for Woodstock College, Maryland," *Monthly Weather Review*, Vol. 87, No. 8, Aug. **1959**, pp. 283–298.

Matalas, N. C., "Some Aspects of Time Series Analysis in Hydrologic Studies:" *Proceedings of the Hydrology Symposium No. 5, McGill University, Montreal, Canada*, McGill University, **1966**, pp. 271–309.

Mitchell, J. Murray Jr., "Climatic Change," World Meteorological Organization *Technical Note* No. 79, Publication No. 195, Geneva, Switzerland, **1966**, 79 pp. (See pp. 33–60.)

Polowchak, Van M., and Panofsky, Hans A., "The Spectrum of Daily Temperatures as a Climatic Indicator," *Monthly Weather Review*, Vol. 96, No. 9, Sept. **1968**, pp. 596–600.

Yule, G. U., "On the Time Series Problem," *Royal Statistical Society Journal*, Vol. 89, **1921**, p. 497.

[*Received August 11, 1972; revised March 2, 1973*]